Available online at www.sciencedirect.com
ScienceDirect
Linear Algebra and its Applications 428 (2008) 382–392
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Orthogonal similarity of a real matrix and its transpose

J. Vermeer *

Delft University of Technology, Delft Institute for Applied Mathematics, P.O. Box 5031, 2600 GA Delft, The Netherlands

Received 24 August 2006; accepted 21 June 2007
Available online 22 August 2007
Submitted by R.A. Horn

Abstract

Any square matrix over a field is similar to its transpose and any square complex matrix is similar to a symmetric complex matrix. We investigate the situation for real orthogonal (respectively complex unitary) similarity.

AMS classification: Primary 15A24

Keywords: Similarity of matrices; Orthogonal similarity of matrices; Transpose

1. Introduction

Let F be a field and let $\mathscr{M}_n(F)$ be the algebra of $n \times n$ matrices. Let $GL_n(F)$ denote the set of invertible $n \times n$ matrices and let $\mathscr{O}_n(F)$ denote the set of orthogonal matrices.

A proof of the following result can be found in [2] (Theorem 4.4.9).

Theorem 1. *Every square complex matrix is similar to a symmetric matrix.*

In this theorem can similarity be replaced by unitary similarity? We will show that this is not the case. In fact we show that for a complex (respectively real) matrix A the property of being

* Tel.: +31 015 2785826; fax: +31 015 2787245.
E-mail address: J.Vermeer@tudelft.nl

0024-3795/$ - see front matter
doi:10.1016/j.laa.2007.06.028

unitarily similar to a symmetric complex matrix is equivalent to the existence of a symmetric unitary (respectively symmetric, real and orthogonal) matrix Q with $QAQ^{\mathrm{T}} = A^{\mathrm{T}}$. Recall the following theorem of Taussky and Zassenhaus.

Theorem 2 [7]. *Let $A \in \mathscr{M}_n(F)$. Then*:

1. *There exists an $X \in GL_n(F)$ such that $XAX^{-1} = A^{\mathrm{T}}$.*
2. *There exists a symmetric $X \in GL_n(F)$ such that $XAX^{-1} = A^{\mathrm{T}}$.*
3. *Every $X \in GL_n(F)$ with $XAX^{-1} = A^{\mathrm{T}}$ is symmetric if and only if the minimal polynomial of A is equal to its characteristic polynomial.*

We restrict ourself to the fields $F = \mathbb{R}$ and $F = \mathbb{C}$.

After constructing examples of real matrices that are not real orthogonally similar to their transposes, the paper [5] came to our attention, in which such an example was constructed. We decided that it is still worthwhile to present our example, since the argument used in [5] is rather deep (it uses Specht's criterion). A question of such an elementary nature deserves an elementary answer.

We present:

(1) Elementary constructions of real matrices A for which no $Q \in \mathscr{O}_n(\mathbb{R})$ exist with $QAQ^{\mathrm{T}} = A^{\mathrm{T}}$. Such a matrix A is not unitarily similar to a symmetric complex matrix.
(2) A matrix $A \in \mathscr{M}_8(\mathbb{R})$ for which a $Q \in \mathscr{O}_8(\mathbb{R})$ exists with $QAQ^{\mathrm{T}} = A^{\mathrm{T}}$, but no symmetric $Q \in \mathscr{O}_8(\mathbb{R})$ exists with this property. Such a matrix A is not unitarily similar to a symmetric complex matrix.
(3) A matrix A for which
  (a) there is some $Q \in \mathscr{O}_n(\mathbb{R})$ such that $QAQ^{\mathrm{T}} = A^{\mathrm{T}}$,
  (b) any such Q is symmetric, and
  (c) the minimal and characteristic polynomial of A do not coincide.

2. The basic results

We start with the following result.

Theorem 3. *Let $A \in \mathscr{M}_n(\mathbb{C})$. The following assertions are equivalent.*

(1) *A is unitarily similar to a complex symmetric matrix.*
(2) *There exists a symmetric unitary matrix U such that UAU^* is symmetric.*
(3) *There exists a symmetric unitary matrix U and a symmetric matrix S such that $A = SU$.*
(4) *There exists a symmetric unitary matrix V such that $VAV^* = A^{\mathrm{T}}$.*

Proof. (1) $\to$ (4) Let O be a unitary matrix and consider OAO^*. Then we have

$$\begin{aligned} &OAO^* \text{ is symmetric} \Leftrightarrow (OAO^*)^{\mathrm{T}} = OAO^* \Leftrightarrow \overline{O}A^{\mathrm{T}}O^{\mathrm{T}} \\ &\quad = OA\overline{O^{\mathrm{T}}} \Leftrightarrow A^{\mathrm{T}} = O^{\mathrm{T}}OA\overline{O^{\mathrm{T}}O}. \end{aligned}$$

Then $V = O^{\mathrm{T}}O$ is the required unitary symmetric matrix.

(4) $\leftrightarrow$ (3) If V is unitary and symmetric then:

AV^* is symmetric $\Leftrightarrow (AV^*)^{\mathrm{T}} = AV^* \Leftrightarrow VAV^* = A^{\mathrm{T}}$.

Hence, put $U = V$ and $S = AV^*$.

(4) $\to$ (2) Being unitary (so normal) and symmetric, V can be written as $V = Q\Lambda Q^{\mathrm{T}}$, with Q real orthogonal and Λ a complex unitary diagonal matrix (see [2], Theorem 4.4.7). Let Δ be a complex unitary diagonal matrix with $\Delta^2 = \Lambda$. Therefore, $U = Q\Delta Q^{\mathrm{T}}$ is unitary and symmetric and $U^{\mathrm{T}}U = U^2 = V$. Then $VAV^* = A^{\mathrm{T}} \Rightarrow U^2 A = A^{\mathrm{T}}U^2 \Rightarrow UA\overline{U} = \overline{U}A^{\mathrm{T}}U = (UA\overline{U})^{\mathrm{T}}$.

(2) $\to$ (1) This case is trivial. $\square$

The following proposition provides us with a large class of examples of matrices that are unitarily similar to a symmetric matrix.

Proposition 4. *Let $A \in \mathscr{M}_n(\mathbb{C})$. If A^2 is normal (in particular, if A is normal), then A is unitarily similar to A^{T} via a symmetric unitary matrix.*

Proof. In [4], Theorem 7.2, it is proved that every $A \in \mathscr{M}_n(\mathbb{C})$ such that A^2 is normal is unitarily similar to a direct sum of blocks, each of which is

$$[\lambda] \quad \text{or} \quad \tau\begin{bmatrix} 0 & 1 \\ \mu & 0 \end{bmatrix}, \quad \tau \in \mathbb{R}, \lambda, \mu \in \mathbb{C}, \tau > 0, \quad \text{and} \quad |\mu| < 1.$$

If we can show that such a block diagonal matrix B is unitarily similar to a complex symmetric matrix, then so is A. According to Theorem 3 it suffices to show that B and B^{T} are similar via a symmetric unitary matrix. If we define $Q = \oplus Q_i$, where $Q_i = [1]$ for each block $[\lambda]$ in B and $Q_i = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ for each block $\tau\begin{bmatrix} 0 & 1 \\ \mu & 0 \end{bmatrix}$ in B, then $Q \in \mathscr{O}_n(\mathbb{R})$, Q is symmetric, and $QBQ^{\mathrm{T}} = B^{\mathrm{T}}$. $\square$

It is known that every $A \in \mathscr{M}_n(\mathbb{C})$ is a product of two symmetric matrices (see [2], Corollary 4.4.11) and part (3) in Theorem 3 presents such a product in which one factor is symmetric unitary. Also it provides a decomposition of A that is quite close to the classical polar decomposition $A = PU$ (where P is positive semidefinite and U unitary). Since in part (2) the symmetry of U can be skipped, it is legitimate to ask the same question in part (4). We will present a real example for which this is not the case. So we need to discuss relations between orthogonal similarity and unitary similarity of real $n \times n$ matrices. One solution can be found in [6]; here is another approach.

Theorem 5. *Let $A \in \mathscr{M}_n(\mathbb{C})$ be nonsingular. Then there are matrices R and E such that:*

(1) *$A = RE$ with $R \in \mathscr{M}_n(\mathbb{R})$, $E\overline{E} = I_n$, and E is polynomial in $\overline{A^{-1}}A$.*
Any such R and E have the following properties:
(2) *$ER = RE$ if and only if $A\overline{A} = \overline{A}A$.*
(3) *If P, Q are real matrices and $P = AQA^{-1}$, then $P = RQR^{-1}$.*
(4) *If A is unitary then R is orthogonal.*
(5) *If A is unitary and symmetric then R is orthogonal and symmetric.*

Proof. The proof of (1) and (2) can be found in [2, Theorem 6.4.23].

(3) Taking the complex conjugate of $P = AQA^{-1}$ we obtain $P = \overline{A}Q\overline{A}^{-1}$ and so we have $P = AQA^{-1} = \overline{A}Q\overline{A}^{-1}$, i.e.,

$$\overline{A}^{-1}AQ = Q\overline{A}^{-1}A.$$

If Q commutes with $\overline{A}^{-1}A$ then it commutes with E, since E is polynomial in $\overline{A}^{-1}A$. Therefore we have

$$P = AQA^{-1} = REQE^{-1}R^{-1} = RQEE^{-1}R^{-1} = RQR^{-1}.$$

(4) If A is unitary then $\overline{A^{-1}}A = A^{\mathrm{T}}A$ is symmetric and hence E (a polynomial in $A^{\mathrm{T}}A$) is also symmetric. The property $E\overline{E} = I_n$ and the symmetry implies that E is unitary. Finally, $R = AE^{-1}$ is unitary and real, so it is orthogonal.

(5) If A is unitary and symmetric then $\overline{A}A = I = A\overline{A}$ and so $A = ER = RE$. It follows that $R = \overline{E}A = A\overline{E}$ and since E (and A) are symmetric we see that $R^{\mathrm{T}} = (\overline{E}A)^{\mathrm{T}} = A^{\mathrm{T}}\overline{E}^{\mathrm{T}} = A\overline{E} = R$, i.e., R is symmetric. □

Corollary 6. *Let $A, B \in \mathscr{M}_n(\mathbb{R})$ be given.*

(1) *A, B are real orthogonally similar if and only if A, B are unitarily similar.*
(2) *A, B are similar via a real symmetric orthogonal matrix if and only if A, B are similar via a symmetric unitary matrix.*

3. Some elementary results

In this section all matrices are assumed to be real. We use the notation Q, Ψ for arbitrary matrices in $\mathscr{O}_n(\mathbb{R})$, Ω for symmetric orthogonal matrices, S for a symmetric matrix, $\mathcal{S}$ for a skew symmetric matrix, D for diagonal matrices and Σ for a real diagonal matrix with $\Sigma^2 = I_n$, that is $\Sigma = \mathrm{diag}(\pm 1, \ldots, \pm 1)$. We say that A has type $\mathcal{S} + D$ if it is the sum of a skew symmetric and a diagonal matrix; we say that A is of type QD if A is the product of an orthogonal and a diagonal matrix. We write $A \simeq B$ (respectively $A \simeq_S B$) if there exists an orthogonal (respectively symmetric and orthogonal) matrix Q with $QAQ^{\mathrm{T}} = B$. The notation $[A]$ denotes the equivalence class of A with respect to $\simeq$. The relation $\simeq_S$ is not an equivalence relation.

The following lemma shows that the properties $A \simeq A^{\mathrm{T}}$ and $A \simeq_S A^{\mathrm{T}}$, and those mentioned in Theorem 11, are properties of $[A]$.

Lemma 7. *Let $A, B \in \mathscr{M}_n(\mathbb{R})$.*

(1) *If $A \simeq A^{\mathrm{T}}$ and $A \simeq B$, then $B \simeq B^{\mathrm{T}}$.*
(2) *If $A \simeq_S A^{\mathrm{T}}$ and $A \simeq B$, then $B \simeq_S B^{\mathrm{T}}$.*
(3) *If $Q \in \mathscr{O}_n(\mathbb{R})$ and $A = QA^{\mathrm{T}}Q^{\mathrm{T}}$, then $A = Q^{\mathrm{T}}A^{\mathrm{T}}Q$.*
(4) *If $A \simeq B$, then $A + A^{\mathrm{T}} \simeq B + B^{\mathrm{T}}$ and $AA^{\mathrm{T}} \simeq BB^{\mathrm{T}}$.*

Proof. (1) and (2) If $A = Q_1BQ_1^{\mathrm{T}}$ and $A = Q_2A^{\mathrm{T}}Q_2$ (Q_1, Q_2 are orthogonal matrices) then $A^{\mathrm{T}} = Q_1B^{\mathrm{T}}Q_1^{\mathrm{T}}$ and so

$$B = Q_1^{\mathrm{T}}AQ_1 = Q_1^{\mathrm{T}}Q_2A^{\mathrm{T}}Q_2Q_1 = Q_1^{\mathrm{T}}Q_2Q_1B^{\mathrm{T}}Q_1^{\mathrm{T}}Q_2Q_1,$$

and we conclude $B \simeq B^{\mathrm{T}}$. If Q_2 is symmetric then $Q_1^{\mathrm{T}}Q_2Q_1$ and so $B \simeq_S B^{\mathrm{T}}$.

(3) $A = QA^{\mathrm{T}}Q^{\mathrm{T}} = QA^{\mathrm{T}}Q^{-1}$ implies $Q^{\mathrm{T}}AQ = A^{\mathrm{T}}$, so $A = (A^{\mathrm{T}})^{\mathrm{T}} = (Q^{\mathrm{T}}AQ)^{\mathrm{T}} = Q^{\mathrm{T}}A^{\mathrm{T}}Q$.

(4) Clear. □

We obtain the real version of Theorem 3.

Theorem 8. *Let $A \in \mathscr{M}_n(\mathbb{R})$. The following assertions are equivalent.*

(1) *A is unitarily similar to a complex symmetric matrix.*
(2) *$A \simeq_S A^{\mathrm{T}}$.*
(3) *$A = \Omega S$, where $\Omega \in \mathscr{O}_n(\mathbb{R})$ is symmetric and S is real symmetric.*
(4) *$A \simeq \Omega' D$, where $\Omega' \in \mathscr{O}_n(\mathbb{R})$ is symmetric and D is real diagonal.*

Proof. (1) $\to$ (2) According to Theorem 3 there exists a unitary symmetric matrix V with $VAV^* = A^{\mathrm{T}}$. Corollary 6 (2) now ensures that $A \simeq_S A^{\mathrm{T}}$.

The equivalence of (1), (2), and (3) can be proved as in Theorem 3.

(3) $\to$ (4) If $A = \Omega S$, then write $S = QDQ^{\mathrm{T}}$ and we obtain

$$A = \Omega QDQ^{\mathrm{T}} \simeq Q^{\mathrm{T}}\Omega QD = \Omega' D.$$

(4) $\to$ (1) If $A \simeq \Omega' D = B$ then $B \simeq_S B^{\mathrm{T}}$ and so by Lemma 7 (2) we have $A \simeq_S A^{\mathrm{T}}$. $\square$

We have two approaches to the problem of constructing a matrix A such that A and A^{T} are not orthogonally similar. The first approach uses the symmetry of $A + A^{\mathrm{T}}$; we show that in the orthogonal similarity equivalence class of A one is free to choose and work with a representative B whose symmetric summand in its Toeplitz decomposition is diagonal. The second approach uses the symmetry of AA^{T}; we show that in the orthogonal similarity equivalence class of A one is free to choose a representative B whose positive semidefinite factor in its polar decomposition is diagonal.

Lemma 9. *Let $A \in \mathscr{M}_n(\mathbb{R})$. Then*:

(1) *$A \simeq B$, where B has type $\mathcal{S} + D$.*
(2) *$A \simeq B$, where B has type QD.*

Proof

(1) Consider the Toeplitz decomposition $A = 1/2(A + A^{\mathrm{T}}) + 1/2(A - A^{\mathrm{T}})$ and suppose $Q(A + A^{\mathrm{T}})Q^{-1} = 2D$. Then $Q^{\mathrm{T}}AQ = D + 1/2Q^{\mathrm{T}}(A - A^{\mathrm{T}})Q$ is the sum of a diagonal matrix and a skew symmetric matrix.
(2) Consider the (real) singular value decomposition $A = Q_1DQ_2^{\mathrm{T}}$. Then $A \simeq B = Q_2^{\mathrm{T}}AQ_2 = \left(Q_2^{\mathrm{T}}Q_1\right)D$. $\square$

For our first approach we also need the following result.

Lemma 10. *Let $A \in \mathscr{M}_n(\mathbb{R})$ be a matrix of type $\mathcal{S} + D$, where $D = \bigoplus_{i=1}^r d_iI_i$ with $d_i \neq d_j$ for $i \neq j$ and I_i the $k_i \times k_i$ identity matrix.*

If $Q \in \mathscr{O}_n(\mathbb{R})$ and $QAQ^{\mathrm{T}} = A^{\mathrm{T}}$ then $Q = \bigoplus_{i=1}^r Q_i$, with $Q_i \in \mathscr{O}_{k_i}(\mathbb{R})$.

Proof. By Lemma 7 (3), $QAQ^{\mathrm{T}} = A^{\mathrm{T}}$ implies $QA^{\mathrm{T}}Q^{\mathrm{T}} = A$ and so $Q(A + A^{\mathrm{T}})Q^{\mathrm{T}} = A + A^{\mathrm{T}}$, i.e., $Q(2D)Q^{\mathrm{T}} = 2D$. Hence, Q commutes with the diagonal matrix $2D$ and this implies the result. (To see this: the first k_1 columns of Q form an orthogonal basis of the eigenspace E_{2d_1} of $2D$ and the standard unit vectors $\{\mathbf{e}_1, \ldots, \mathbf{e}_{k_1}\}$ (in $\mathbb{R}^n$) is a basis of E_{2d_1}.) $\square$

Theorem 11. *Assume that $A \in \mathscr{M}_n(\mathbb{R})$ satisfies one of the following properties*:

(1) *A has n different real eigenvalues*, *or*
(2) *$A + A^{\mathrm{T}}$ has n different eigenvalues*, *or*
(3) *A has n different singular values*,

If there exists $Q \in \mathscr{O}_n(\mathbb{R})$ with $A = QA^{\mathrm{T}}Q^{\mathrm{T}}$, then Q is symmetric and there exist at most 2^n such matrices.

Proof. (1) The symmetry of Q is a direct corollary of Theorem 2. We only have to check that there exist at most 2^n such matrices Q.

Let $D \in \mathscr{M}_n(\mathbb{R})$ be a diagonal matrix whose diagonal entries are the (distinct) eigenvalues of A. Suppose $P, R \in \mathscr{M}_n(\mathbb{R})$ are nonsingular and diagonalize A, i.e., $A = PDP^{-1} = RDR^{-1}$. Corresponding columns of P and R are nonzero vectors in the same one-dimensional eigenspace of A, so there is a diagonal matrix $C \in \mathscr{M}_n(\mathbb{R})$ such that $R = PC$. We fix one specific matrix $P \in \mathscr{M}_n(\mathbb{R})$ with $A = PDP^{-1}$.

Now suppose $Q, \Psi \in \mathscr{O}_n(\mathbb{R})$ are such that $A = QA^{\mathrm{T}}Q^{\mathrm{T}} = \Psi A^{\mathrm{T}}\Psi$, so

$$A = Q(PDP^{-1})^{\mathrm{T}}Q^{\mathrm{T}} = (Q(P^{\mathrm{T}})^{-1})D(Q(P^{\mathrm{T}})^{-1})^{-1}$$

and

$$A = \Psi(PDP^{-1})^{\mathrm{T}}\Psi^{\mathrm{T}} = (\Psi(P^{\mathrm{T}})^{-1})D(\Psi(P^{\mathrm{T}})^{-1})^{-1}.$$

Note that $Q(P^{\mathrm{T}})^{-1} = PC_1$, for some diagonal matrix $C_1 \in \mathscr{M}_n(\mathbb{R})$, so $Q = PC_1P^{\mathrm{T}}$ is symmetric. Next we note that $Q(P^{\mathrm{T}})^{-1} = \Psi(P^{\mathrm{T}})^{-1}C_2$ for some diagonal matrix $C_2 \in \mathscr{M}_n(\mathbb{R})$, so

$$P^{\mathrm{T}}(\Psi^{-1}Q)(P^{\mathrm{T}})^{-1} = C_2$$

and we have obtained a real diagonalization of the real orthogonal matrix $\Psi^{-1}Q$, whose only possible (real) eigenvalues are ± 1. Thus $C_2 = \mathrm{diag}(\pm 1, \ldots, \pm 1)$ which shows that there are at most 2^n choices for Ψ, since $\Psi = Q(P^{\mathrm{T}})^{-1}C_2^{-1}P^{\mathrm{T}}$.

(2) If $Q \in \mathscr{O}_n(\mathbb{R})$ and $A = Q^{\mathrm{T}}A^{\mathrm{T}}Q$ then, by Lemma 7 (3), $A^{\mathrm{T}} = QAQ^{\mathrm{T}}$ and so $Q(A + A^{\mathrm{T}})Q^{\mathrm{T}} = A + A^{\mathrm{T}} = (A + A^{\mathrm{T}})^{\mathrm{T}}$; the conclusions follow from part (1).

(3) Fix $Q \in \mathscr{O}_n(\mathbb{R})$ with the property $A = QA^{\mathrm{T}}Q^{\mathrm{T}}$ and consider a singular value decomposition $A = UDV^{\mathrm{T}}$ of A. Then $A = QA^{\mathrm{T}}Q^{\mathrm{T}} = (QV)D(QU)^{\mathrm{T}}$, so we have two singular value decompositions of A. Distinct singular values and the uniqueness theorem for the singular value decomposition (see [3], Theorem 3.1.1) ensure that $QV = UP$ and $QU = VP$ for some diagonal real orthogonal matrix P. Then $Q = UPV^{\mathrm{T}} = VPU^{\mathrm{T}} = Q^{\mathrm{T}}$ is symmetric and there are only 2^n choices for P. $\square$

Remark 12. Let $A \in \mathscr{M}_n(\mathbb{R})$ with n different (possibly complex) eigenvalues. If there exists $Q \in \mathscr{O}_n(\mathbb{R})$ with $A = QA^{\mathrm{T}}Q^{\mathrm{T}}$, then the symmetry of Q is still a direct corollary of Theorem 2. However, Example 5 shows that there might exist infinitely many of such Q if not all eigenvalues are real.

We end this section with some positive results.

Proposition 13

(1) *Let $A \in \mathscr{M}_2(\mathbb{R})$. Then $A \simeq_S A^{\mathrm{T}}$.*
(2) *Let $A \in \mathscr{M}_3(\mathbb{R})$. If $A \simeq A^{\mathrm{T}}$ then $A \simeq_S A^{\mathrm{T}}$.*

Proof

(1) Lemma 9 implies that there exists a matrix B of type $\mathcal{S}+D$ with $A \simeq B$. If we can show that $B \simeq_S B^{\mathrm{T}}$ then Lemma 7 (2) implies that $A \simeq_S A^{\mathrm{T}}$. Since $B = \begin{bmatrix} p & r \\ -r & q \end{bmatrix}$, it is clear that the orthogonal symmetric matrix $\Omega = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \in \mathcal{O}_2(\mathbb{R})$ satisfies $\Omega B \Omega^{\mathrm{T}} = B^{\mathrm{T}}$.

(2) Lemma 9 implies that there exists a matrix B of type $\mathcal{S}+D$ with $A \simeq B$ and Lemma 7 (1) implies $B \simeq B^{\mathrm{T}}$. If we can show that $B \simeq_S B^{\mathrm{T}}$, then Lemma 7 (2) implies that $A \simeq_S A^{\mathrm{T}}$. Write $B = \mathcal{S} + D$, where $\mathcal{S}$ is skew symmetric and D is diagonal.

Case 1. If the three diagonal entries of D are different then the assertion $B \simeq_S B^{\mathrm{T}}$ follows from Theorem 11.

Case 2. If the three diagonal entries of D are identical, then $B = \mathcal{S} + dI_3$. The matrix $\mathcal{S}$ is skew symmetric and real so $\mathcal{S}$ is normal. By Proposition 4 and Corollary 6 there exists a real symmetric orthogonal matrix Ω with $\Omega\mathcal{S}\Omega = \mathcal{S}^{\mathrm{T}}$. We conclude that $\Omega(\mathcal{S}+dI_3)\Omega = (\mathcal{S}+dI_3)^{\mathrm{T}}$, i.e., $B \simeq_S B^{\mathrm{T}}$.

Case 3. There are two values among the three entries of D.

Write $B = \begin{bmatrix} 0 & s_1 & s_2 \\ -s_1 & 0 & s_3 \\ -s_2 & -s_3 & 0 \end{bmatrix} + \begin{bmatrix} d_1 & 0 & 0 \\ 0 & d_1 & 0 \\ 0 & 0 & d_2 \end{bmatrix}$. From Lemma 10 any $Q \in \mathcal{O}_3(\mathbb{R})$ with $QBQ^{\mathrm{T}} = B^{\mathrm{T}}$ has the structure: $Q = \begin{bmatrix} Q_1 & 0 \\ 0 & \pm 1 \end{bmatrix}$ with $Q_1 \in \mathcal{O}_2(\mathbb{R})$. Put $B_1 = \begin{bmatrix} d_1 & s_1 \\ -s_1 & d_1 \end{bmatrix}$ and $x = \begin{bmatrix} s_2 \\ s_3 \end{bmatrix}$.

Assume $x = 0$. The first part of this Proposition shows that there is a symmetric Ω_1 with $\Omega_1 B_1 \Omega_1 = B_1^{\mathrm{T}}$. Put $\Omega = \begin{bmatrix} \Omega_1 & 0 \\ 0 & \pm 1 \end{bmatrix}$ and since $B = \begin{bmatrix} B_1 & 0 \\ 0 & d_3 \end{bmatrix}$ we obtain $\Omega B \Omega = B^{\mathrm{T}}$.

Assume $x \neq 0$. Since $QBQ^{\mathrm{T}} = B^{\mathrm{T}}$ we see

$$\begin{bmatrix} Q_1 & 0 \\ 0 & \pm 1 \end{bmatrix} \begin{bmatrix} B_1 & x \\ -x^{\mathrm{T}} & d_3 \end{bmatrix} \begin{bmatrix} Q_1^{\mathrm{T}} & 0 \\ 0 & \pm 1 \end{bmatrix} = \begin{bmatrix} B_1^{\mathrm{T}} & -x \\ x^{\mathrm{T}} & d_3 \end{bmatrix}$$
$$\Rightarrow \begin{bmatrix} Q_1 B_1 Q_1^{\mathrm{T}} & \pm Q_1 x \\ \mp x^{\mathrm{T}} Q_1^{\mathrm{T}} & d_3 \end{bmatrix} = \begin{bmatrix} B_1^{\mathrm{T}} & -x \\ x^{\mathrm{T}} & d_3 \end{bmatrix}$$

We conclude that x is a real eigenvector of Q_1. A 2×2 real orthogonal matrix is either of type $\begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$ with complex eigenvalues $\mathrm{e}^{\pm \mathrm{i}\varphi}$ and associated nonreal eigenvectors, or of type $\begin{bmatrix} \cos\varphi & \sin\varphi \\ \sin\varphi & -\cos\varphi \end{bmatrix}$ with eigenvalues ± 1. So a 2×2 real orthogonal matrix with a real eigenvector is symmetric, i.e., Q_1 (and Q) must be symmetric. The conclusion $B \simeq_S B^{\mathrm{T}}$ follows. $\square$

4. The examples

As promised in the Introduction we use two elementary methods to construct real matrices A for which there is no $Q \in \mathcal{O}_n(\mathbb{R})$ such that $QAQ^{\tau} = A^{\tau}$.

Method 1. If $A \in \mathcal{M}_n(\mathbb{R})$ has type $\mathcal{S}+D$ $(n > 2)$ and D has n different diagonal entries and no entry of $\mathcal{S}$ outside the diagonal is 0, then there is no $Q \in \mathcal{O}_n(\mathbb{R})$ with $QAQ^{\tau} = A^{\tau}$.

Proof. Lemma 10 implies that any Q with $QAQ^{\mathrm{T}} = A^{\mathrm{T}}$ must be of type Σ. Since $n > 2$, Σ will have two columns with the same entry on the diagonal, say column i and j. It is easy to verify that $[\Sigma A\Sigma]_{i,j} = [A]_{i,j}$. Since no entry of A outside the diagonal is 0 we conclude $\Sigma A\Sigma \neq A^{\mathrm{T}}$ for all Σ. $\square$

Example 1. Consider $A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & -1 \end{bmatrix}$. Method 1 shows that A is not orthogonally similar to A^{T}. This example can be used in any field $\mathbb{F}$ of characteristic $\neq 2$ to obtain an $A \in \mathscr{M}_n(\mathbb{F})$ for which no $Q \in \mathscr{O}_n(\mathbb{F})$ exists with $QAQ^{\tau} = A^{\tau}$.

Lemma 14. *Let $A \in \mathscr{M}_n(\mathbb{R})$ have n different singular values. If $A \simeq QD$, with $Q \in \mathscr{O}_n(\mathbb{R})$ and D a diagonal matrix, then the following assertions are equivalent.*

(1) $A \simeq A^{\mathrm{T}}$.
(2) *There exists a diagonal matrix $\Sigma = \mathrm{diag}(\pm 1, \ldots, \pm 1)$ such that $Q\Sigma$ is symmetric.*

Proof. (1) $\Rightarrow$ (2). Consider $B = QD$. Then B has n different singular values, as $B \simeq A$. And we also obtain that $B \simeq B^{\mathrm{T}}$. By Theorem 11 any real orthogonal similarity of B to B^{T} must be symmetric, so $B \simeq_S B^{\mathrm{T}}$. Fix a real, symmetric, and orthogonal matrix Ω with $B = \Omega B^{\mathrm{T}}\Omega$.

In the proof of Theorem 11 (3) we showed: if $B = UDV^{\mathrm{T}}$ is a singular value decomposition then there exists a real orthogonal matrix P of type Σ with $\Omega V = UP$.

A singular value decomposition of B can be obtained by $V = I$, $D_1 = \Sigma_1 D$ (with $[\Sigma_1]_{i,i} = \mathrm{signum}[D]_{i,i}$) and $U = Q\Sigma_1$. So there exists a matrix $P = \Sigma_2$ with $\Omega = \Omega I = Q\Sigma_1\Sigma_2$, i.e., $\Sigma_1\Sigma_2$ is of type Σ and $Q\Sigma_1\Sigma_2$ is symmetric.

(2) $\Rightarrow$ (1) $A \simeq QD = Q\Sigma^2 D = (Q\Sigma)(\Sigma D) = \Omega D'$ and use Theorem 8. $\square$

Method 2. Let $A = Q_1 D \in \mathscr{M}_n(\mathbb{R})$ $(n > 2)$ in which $Q_1 = [q_{i,j}] \in \mathscr{O}_n(\mathbb{R})$ and $D \in \mathscr{M}_n(\mathbb{R})$, a diagonal matrix with n different nonnegative diagonal entries. If there exist distinct i, j such that $|q_{i,j}| \neq |q_{j,i}|$, then there is no $Q \in \mathscr{O}_n(\mathbb{R})$ with $QAQ^{\tau} = A^{\tau}$.

Proof. The diagonal entries of D are the n different singular values of A. If A is orthogonally similar to A^{T}, then $Q_1\Sigma$ is symmetric for some real diagonal matrix with the property $\Sigma^2 = I$. It follows that $|q_{i,j}| = |q_{j,i}|$, for all i, j. $\square$

Example 2 (*This example is due to P. Sonneveld who had a numerical explanation*)

$$A = \begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & 1 & -2 \end{bmatrix} = QD = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \\ -1/\sqrt{2} & 1/\sqrt{3} & 1/\sqrt{6} \\ 0 & 1/\sqrt{3} & -2/\sqrt{6} \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & \sqrt{6} \end{bmatrix}.$$

Method 2 shows that A is not orthogonally similar to A^{T}.

We now present an example of a matrix with the property $A \simeq A^{\mathrm{T}}$ but not $A \simeq_S A^{\mathrm{T}}$.

Example 3. Consider the following 8×8 matrix:

$$A = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & -1 & 0 & -1 & 0 & -1 \\ -1 & 0 & 2 & 0 & 1 & 1 & 2 & 1 \\ 0 & 1 & 0 & 2 & 1 & -1 & 1 & -2 \\ -1 & 0 & -1 & -1 & 3 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 & 0 & 3 & 0 & -1 \\ -1 & 0 & -2 & -1 & -1 & 0 & 4 & 0 \\ 0 & 1 & -1 & 2 & 0 & 1 & 0 & 4 \end{bmatrix}$$

which we represent as a 4×4 block matrix $\{A_{i,j}\}_{i \leqslant 4, j \leqslant 4}$ of 2×2 blocks $A_{i,j}$, with

$$A_{i,i} = \begin{bmatrix} i & 0 \\ 0 & i \end{bmatrix} \text{ for } i = 1, \dots, 4,$$

$$A_{1,2} = A_{1,3} = A_{1,4} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = A_{3,4},\ A_{2,3} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad \text{and} \quad A_{2,4} = \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}.$$

Finally, for $i > j$ we define $A_{i,j} = -A^{\mathrm{T}}_{j,i} (= -A_{j,i})$. We claim that

(1) $A \simeq A^{\mathrm{T}}$ but not $A \simeq_S A^{\mathrm{T}}$.
(2) A is not unitarily similar to a symmetric (complex) matrix.
(3) A is not (complex) orthogonally similar to a symmetric (complex) matrix.

Proof. (1) We check $A \simeq A^{\mathrm{T}}$.

(a) A is of type $\mathcal{S} + D$.
(b) Lemma 10 implies that any $Q \in \mathcal{O}_8(\mathbb{R})$ with $QAQ^{\mathrm{T}} = A^{\mathrm{T}}$ is block diagonal: $Q = \bigoplus_{i=1}^4 Q_i$, with $Q_i \in \mathcal{O}_2(\mathbb{R})$.
(c) If $Q = \bigoplus_{i=1}^4 Q_i \in \mathcal{O}_8(\mathbb{R})$ (with $Q_i \in \mathcal{O}_2(\mathbb{R})$) then QAQ^{T} is the block matrix $\{Q_i A_{i,j} Q_j^{\mathrm{T}}\}_{i \leqslant 4, j \leqslant 4}$.
(d) Therefore, $Q = \bigoplus_{i=1}^4 Q_i \in \mathcal{O}_8(\mathbb{R})$ satisfies $QAQ^{\mathrm{T}} = A^{\mathrm{T}}$ if and only if $Q_i A_{i,i} {Q_i}^{\mathrm{T}} = A_{i,i}$ for each i (satisfied for any $Q_i \in \mathcal{O}_2(\mathbb{R})$ since each $A_{i,i}$ is a scalar matrix) and $Q_i A_{i,j} Q_j^{\mathrm{T}} = -A_{i,j}$ for $i \neq j$.
The latter conditions are satisfied for $Q_1 = Q_2 = Q_3 = Q_4 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

If there were a symmetric $\Omega \in \mathcal{O}_8(\mathbb{R})$ such that $\Omega A \Omega = A^{\mathrm{T}}$, then $\Omega = \bigoplus_{i=1}^4 \Omega_i$ with symmetric $\Omega_i \in \mathcal{O}_2(\mathbb{R})$. In this case:

(a) $\Omega_1 A_{1,2} \Omega_2 = -A_{1,2}$, $\Omega_1 A_{1,3} \Omega_3 = -A_{1,3}$, and $\Omega_1 A_{1,4} \Omega_4 = -A_{1,4}$. Since $A_{1,2} = A_{1,3} = A_{1,4}$ (is invertible) this implies that $\Omega_2 = \Omega_3 = \Omega_4$.
(b) So $\Omega_2 (= \Omega_3)$ satisfies: $\Omega_2 A_{2,3} \Omega_2 = -A_{2,3}$ and one checks that $\Omega_2 = \pm \begin{bmatrix} -\sqrt{2}/2 & \sqrt{2}/2 \\ \sqrt{2}/2 & \sqrt{2}/2 \end{bmatrix}$ are the only two real, symmetric, and orthogonal matrices with $\Omega_2 A_{2,3} \Omega_2 = -A_{2,3}$.
(c) We obtain a contradiction by observing that $\Omega_2 A_{2,4} \Omega_4 = \begin{bmatrix} -1 & -2 \\ -2 & 1 \end{bmatrix} \neq -A_{2,4}$.

(2) Since this matrix A does not have the property $A \simeq_S A^{\mathrm{T}}$, Theorem 8 implies that A is not unitarily similar to a complex symmetric matrix.

(3) If $A = QBQ^{\mathrm{T}}$, with Q (complex) orthogonal and B a (complex) symmetric matrix then A must be symmetric. $\square$

Next we present an example of a matrix $A \in \mathcal{M}_3(\mathbb{R})$ for which

(1) there is some $Q \in \mathcal{O}_3(\mathbb{R})$ such that $QAQ^{\mathrm{T}} = A^{\mathrm{T}}$,
(2) any such Q is symmetric, and
(3) the minimal and characteristic polynomial of A do not coincide.

Example 4. Consider $A = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. Note that $QAQ^{\mathrm{T}} = A^{\mathrm{T}}$ if $Q = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ and that $A + A^{\mathrm{T}}$ has three different eigenvalues. We conclude that any $Q \in \mathscr{O}_3(\mathbb{R})$ such that $QAQ^{\mathrm{T}} = A^{\mathrm{T}}$ is symmetric and A satisfies the three required properties.

One can also check that any unitary U with $UAU^* = A^{\mathrm{T}}$ is symmetric. For such a U it follows that $UA^{\mathrm{T}}U^* = A$, so $U(A + A^{\mathrm{T}})U^* = A + A^{\mathrm{T}} = (A + A^{\mathrm{T}})^{\mathrm{T}}$; use Theorem 2 (3) with the matrix $A + A^{\mathrm{T}}$. The minimal and characteristic polynomial of $A + A^{\mathrm{T}}$ coincide.

The same argument shows that any (complex) orthogonal U with $UAU^{\mathrm{T}} = A^{\mathrm{T}}$ is symmetric.

Next we present an example of a matrix $A \in \mathscr{M}_2(\mathbb{R})$ with two different complex eigenvalues for which infinitely many $Q \in \mathscr{O}_2(\mathbb{R})$ exists such that $QAQ^{\mathrm{T}} = A^{\mathrm{T}}$.

Example 5. Consider $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. The matrix A has two different eigenvalues: $\pm i$. However, every symmetric $Q \in \mathscr{O}_2(\mathbb{R})$ has the property: $QAQ^{\mathrm{T}} = A^{\mathrm{T}}$. See Theorem 11 (1) and Remark 12.

Theorem 2 (1) fails if *field* is replaced by *ring*; see [1] (text after Theorem 6) for an example in $GL(2, \mathbb{Z})$. The following example shows that it also fails over a skew field. The final example shows that in an real infinite dimensional Hilbert space an operator need not be similar to its adjoint (transpose).

Example 6 (*H. Kneppers and I noticed the following*). Consider the real quaternions $\mathbb{H}$, the linear associative algebra over the real numbers having as a basis four independent elements $1, i, j, k$ with $i^2 = j^2 = k^2 = ijk = -1$. It follows that $ij = -ji = k$, $jk = -kj = i$ and $ki = -ik = j$. The matrix

$$K = \begin{bmatrix} 1 & i \\ j & k \end{bmatrix}$$

is invertible and $K^{-1} = \begin{bmatrix} 1/2 & -j/2 \\ -i/2 & -k/2 \end{bmatrix}$.

However, we claim that $K^{\mathrm{T}} = \begin{bmatrix} 1 & j \\ i & k \end{bmatrix}$ is not invertible. For, assume

$$\begin{bmatrix} 1 & j \\ i & k \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Entry (1, 1) implies: $a + jc = 1 \Rightarrow i(a + jc) = i \Rightarrow ia + kc = i$.

Entry (2, 1) implies: $ia + kc = 0$.

We have obtained a contradiction and we have proved the claim. But this implies that K and K^{T} cannot be similar, as invertibility is preserved by similarity.

Example 7. Consider the real Hilbert space ℓ_2 and the shift operator $\sigma : \ell_2 \to \ell_2$,

$$\sigma(x_1, \ldots, x_n, \ldots) = (x_2, \ldots, x_n, \ldots).$$

The transpose (or adjoint) of σ is defined by the formula:

$$\sigma^{\mathrm{T}}(x_1, \ldots, x_n, \ldots) = (0, x_1, \ldots, x_n, \ldots).$$

The operator σ is surjective but not injective and σ^{T} is injective but not surjective. So σ and σ^{T} cannot be similar, i.e., there exists no operator $\tau : \ell_2 \to \ell_2$ such that $\tau\sigma\tau^{-1} = \sigma^{\mathrm{T}}$.

References

[1] J.P. Henniger, Factorization and similarity in $GL(2, \mathbb{Z})$, Linear and Multilinear Algebra 251 (1997) 223–237.
[2] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, 1985.
[3] R.A. Horn, C.R. Johnson, Topics in Matrix Analysis, Cambridge University Press, 1991.
[4] R.A. Horn, V. Sergeichuk, Canonical forms for unitary congruence and *congruence, Linear and Multilinear Algebra, in press.
[5] Kh.D. Ikramov, On the unitary similarity of mutually transposed matrices (Russian), Dokl. Akad. Nauk 378 (2001) 592–593;
On the unitary similarity of a matrix and its transpose, Dokl. Math. 63 (2001) 377–378.
[6] D.I. Merino, Solution, IMAGE (Bulletin of the International Linear Algebra Society) 21 (1998) 26–27.
[7] O. Taussky, H. Zassenhaus, On the similarity transformation between a matrix and its transpose, Pacific J. Math. 9 (1959) 893–896.

Further readings

[1] J.W. Duke, A note on the similarity of a matrix and its conjugate transpose, Pacific J. Math. 31 (1969) 321–323.
[2] R. Gow, The equivalence of an invertible matrix to its transpose, Linear and Multilinear Algebra 8 (1980) 371–373.
[3] R.A. Horn, D.I. Merino, A real-coninvolutory analog of the polar decomposition, Linear Algebra Appl. 190 (1993) 209–227.